# Multiple Baseline Single-Case Experimental Design Using Statistical Analysis: Piloting a Mindfulness-Based Intervention for Obsessive–Compulsive Disorder


## Megan Wilkinson-Tough

University of Bath, UK


## Discipline

Psychology [D3]

## Sub-discipline

Clinical and Abnormal Psychology [SD-Psych-3]

## Academic Level

Postgraduate

## Contributor Biography

**Dr. Megan Wilkinson-Tough** is a clinical research tutor for the Bath Doctorate in Clinical Psychology programme. She qualified as a clinical psychologist from the Bristol Doctorate in

Clinical Psychology in 2008 and completed a diploma in Schema Focused Therapy in 2013. She is a chartered clinical psychologist with the British Psychological Society and is registered by the Health and Care Professions Council. Since qualification, she has specialized in working in adult mental health settings, including secure settings. She currently works clinically within the Centre for Specialist Psychological Treatments of Anxiety and Related Problems.

## Published Articles

Wilkinson-Tough, M., Bocci, L., Thorne, K., & Herlihy, J. (2010). Is mindfulness-based therapy an effective intervention for obsessive–intrusive thoughts: A case series. *Clinical Psychology & Psychotherapy, 17*, 250-268.

## Abstract

Single-case experimental designs provide an ideal research method for the initial stages of testing new interventions. As part of my Doctorate in Clinical Psychology, I sought to test a mindfulness-based approach to intrusive thoughts associated with obsessive–compulsive disorder. At this time, the evidence for mindfulness-based interventions was growing but a purely mindfulness-based approach to psychological therapy had not been trialed in obsessive–compulsive disorder despite there being theoretical reasons to believe that it might be helpful for this often debilitating mental health problem. Single-case experimental designs allow you to gather rich detail about the therapy process, providing a deeper understanding about how an intervention works for individuals when little is known about the treatment. Repeatedly demonstrating the effect of moving from a baseline (no intervention) to a treatment phase in a multiple baseline design helps increase confidence that any effect found is produced by the intervention rather than a co-occurring factor. In this case study, I appraise the use of a multiple

baseline design in which three participants moved from baseline monitoring of obsessive–compulsive disorder symptoms, to a control intervention (relaxation training), and to a mindfulness-based intervention for obsessive–compulsive disorder. I discuss the advantages of using both standardized and individually tailored (idiographic) measures of outcomes. I also consider the arguments for and against the use of statistical analysis with the type of data collected in single-case experimental designs and discuss the challenges encountered in undertaking this type of analysis.

## Learning Outcomes

By the end of this case, students should be able to

• Understand and describe the basic concepts of single-subject research and its application within the field of mental health research

• Understand how the psychological processes of Thought-Action Fusion and Thought Suppression are thought to contribute to obsessive–compulsive symptoms and why a mindfulness intervention could help with this

• Understand the advantages and disadvantages of using standardized versus idiographic measures

• Be able to weigh up the value of statistical analysis versus visual inspection of time-series data

## Case Study

# Project Overview and Context

Obsessive–compulsive disorder (OCD) is a distressing mental health problem that can occupy much of the time of those who experience it with obsessive, anxiety-provoking intrusive thoughts

and compulsive rituals (American Psychiatric Association [APA], 2013). Clinicians and researchers continue to work to improve treatments for OCD to increase recovery rates and reduce the likelihood that symptoms will return post-treatment. In recent years, research has looked to new approaches which might complement cognitive-behavior therapy (CBT), an evidence-based treatment for OCD. Mindfulness meditation practices were first used within psychological therapies in the 1990s and have been integrated into CBT for other mental health problems. I had developed an interest in this area having seen the effectiveness of these approaches in my work as an Assistant Psychologist before I commenced my Clinical Psychology doctoral training in 2005. At the time no one had researched the effectiveness of using mindfulness as a stand-alone treatment to improve OCD symptoms despite there being good theoretical reasons as to why it might be helpful. I decided that SCED methodology would allow me to pilot a mindfulness-based intervention for OCD, providing "proof of concept" (an initial demonstration of efficacy) for this therapeutic approach.

## Underpinning Theory

People who have OCD experience intrusive thoughts that distress them (e.g., "I might have left the gas on") and they try and get rid of (neutralize) the distress caused by these thoughts by performing compulsive actions (e.g., repeatedly checking the gas cooker). When trying to understand how OCD develops, Rachman and de Silva (1978) asked people who did and did not have OCD about their intrusive thoughts and found they were very similar in content. Very many people experience thoughts that pop into their mind, some of which may seem unpleasant (e.g., "I'm going to jump in front of that train" or "I'm going to attack someone I love") but they don't cause most people problems. What seems to be different about people who develop OCD is that

they are likely to believe that unwanted intrusive thoughts are important, whereas people who do not have OCD can easily dismiss them. People with OCD often believe that they might act on an intrusive thought, are responsible for stopping something they thought from happening, or are morally bad simply for having the thought. Stanley Rachman (1993) called this process psychological "fusion," later known as Thought-Action Fusion (TAF).

Because people with OCD find such intrusive thoughts worrying they try to get rid of them by suppressing or neutralizing them, but research has found that the more effort people put into trying to get rid of thoughts, the more such thoughts return (Wenzlaff & Wegner's, 2000, review of Thought Suppression is good if you are interested in reading further about this). These ideas form the basis of a cognitive-behavioral model of OCD which outlines how believing that intrusive thoughts are significant and trying to suppress/neutralize them to avoid feeling distressed sets up a cycle where obsessions and compulsions develop.

To me, mindfulness techniques seem to offer an alternative to the cycle of TAF and Thought Suppression that is thought to contribute to OCD symptoms. Mindfulness meditation is drawn from Buddhist practice but can also be used by people who are not religious. Mindfulness practice emphasizes the importance of seeing thoughts as "just thoughts" and of accepting (rather than suppressing) all types of experience (including thoughts), even if we might find such experiences difficult. I thought that this approach could complement CBT, which is the recommended first-line psychological treatment for OCD (NICE, 2005).

Although there is clear evidence that CBT is the most effective psychological intervention available for OCD (Gava et al., 2007), some people do not fully benefit from this approach; some people find the therapy tasks too difficult, whereas others find that CBT helps

only partially or that their symptoms return after therapy (e.g., Eisen et al., 1999). There was a need to investigate potential complementary approaches such as mindfulness, and I was keen to contribute to this process. However, there are ethical considerations when researching a new therapy application where the effects are not known. In such circumstances, it is best to initially trial this on a very small number of people who are made aware of the experimental nature of the treatment. You can then collect in-depth information about the effects and experience of the treatment for each individual before testing it on larger groups.

## Research Design

SCEDs allow you to study individual treatment responses in a methodologically rigorous way. Repeated measurement of the problem (the Dependent Variable [DV]) before an intervention allows you to establish the baseline level of difficulty for that individual. A single measurement of the problem does not allow you to do this because it only provides a one-off snapshot of a difficulty which might change (and indeed might improve on its own without intervention). Repeated measurement of a problem that does not change over a set time period (known as a stable baseline) provides evidence that the DV—in this case distress caused by OCD symptoms, TAF, and Thought Suppression—would remain the same without intervention. Once a stable baseline is established (known as Phase A), you can then move to Phase B, the introduction of the intervention (in research terms this is the manipulation of the Independent Variable [IV]). If a difference in repeated measurement of the DV is found after the introduction of the intervention, then you have evidence that the intervention had an effect.

The evidence for the effect of the intervention is strengthened by repeatedly moving from a baseline to an intervention and demonstrating the effect of the intervention each time. In certain

circumstances, you can demonstrate this with a single individual. For example, you could stop the intervention and go back to baseline monitoring and then reintroduce the intervention (known as a reversal or ABAB design, where A is the baseline phase and B is the intervention). This only works if the effects of the intervention are reversible should you stop treatment, which is infrequently the case with psychological therapies. You could also systematically apply the intervention to different aspects of a person's problem, one at a time, and if the intervention is effective it will sequentially affect each aspect of the problem (e.g., anxiety at home, anxiety at school, anxiety in social activities) as the intervention is introduced; this is known as a within-participant multiple baseline design.

Neither of these designs perfectly fit my intended research. I thought that once some of the ideas involved in the mindfulness intervention were understood by participants, they could not then be unlearnt (i.e., the intervention could not be reversed). Equally, I thought it would be hard to apply mindfulness learning in a way that selectively targeted just one aspect of the person's problem. Because of this I chose an alternative across-participant multiple baseline design, where you monitor the baseline difficulties of a number of different participants before introducing an intervention. Repeatedly showing the effect of moving from baseline to intervention across a number of participants strengthens the evidence that it is the intervention, rather than another factor (e.g., a change in the weather and natural recovery), that produced the improvement.

I was also aware that simply coming to therapy sessions and participating in an activity that participants thought might be beneficial could produce an effect in itself. Because of this I decided to include a control or placebo "intervention" to help strengthen the evidence that any positive effects of the intervention were due to the mindfulness itself. I chose to use a phase of

relaxation training as a control condition, as had been done in previous OCD research (e.g., Greist, Marks, & Baer, 2002), because the training could be made to look very much like the mindfulness intervention (e.g., practice in sessions and daily practice guided by a recording). I therefore chose an A (Baseline Monitoring) B (Relaxation Control) C (Mindfulness Intervention) design.

## Design Practicalities

I had to make decisions about what aspects of the design I wanted to prioritize. I thought that it was important to ensure that I had a way of confirming whether the DV (TAF, Thought Suppression and Distress) were stable or whether they were changing. I chose to use the C Statistic developed by Young in 1941 (detailed in Equation 1) as a means of analyzing this

$$C = 1 - \frac{\sum_{i=1}^{N-1} (Xi - Xi+1)^2}{2\sum_{i=1}^{N} (Xi - X)^2}$$

(1)

Equation 1 defines The C Statistic (Young, 1941)

The C statistic provides a technique to test whether the time-series data often produced by SCED (i.e., repeated measurements taken at equal, regular intervals, such as daily or weekly) is stable or whether it follows an increasing or decreasing trend. For example, when testing an intervention for OCD you would predict that OCD symptoms remain stable before you start therapy but would improve (show a decreasing trend, or downward slope, on a graph symptom severity) once an effective therapy started. You need a minimum of eight measurements in each phase to reliably use this statistical technique (Young, 1941). I thought that it would be difficult and ethically problematic to take weekly measurements, as this would mean keeping participants

in both the phases of baseline monitoring and the relaxation control for a minimum of 8 weeks each, which was simply too long. I therefore decided that measurements would need to be taken daily. However, this ruled out the possibility of using lengthy standardized measures on a daily basis because generally such measures are not designed to be used with such frequency and they are usually too long to feasibly ask participants to complete daily. I therefore decided to develop my own brief measures that would capture the individual nature of participants' difficulties (known as idiographic measurement) alongside less frequent use of standardized questionnaires.

The daily ratings system I developed to measure TAF required participants to rate on a scale of 0 to 100 how much they believed their interpretation of their intrusive thoughts; for example, "Having a thought of stabbing my husband means that I might lose control and actually stab him." The wording of this scale was developed with each participant in an assessment session to capture their specific concerns. I also asked them to rate on the same 0 to 100 scale how much they attempted to "suppress your obsessive thought and push it completely from your mind," which I used as a measure of Thought Suppression. They additionally rated the amount of distress they felt when experiencing the thought on the same 0 to 100 scale. Although none of these scales had been through robust psychometric development, similar versions had been used in previous peer-reviewed research (e.g., Masuda, Hayes, Sackett, & Twohig, 2004). The benefit of the rating scales I developed was that they took under a minute a day for participants to complete and captured the difficulties they experienced in their own words. However, factors that could increase measurement error with such measures should be borne in mind, such as the possibility that participant's increasing familiarity with and memory for the items with repeated measurement changed the way that they responded to the questions over time.

Due to the increased potential for measure error using idiographic measures, it was important to cross-reference (or triangulate) the ratings developed for this research with standardized tools, to provide some evidence for the validity of these new measures. I therefore used the Yale-Brown Obsessive-Compulsive Scale (YBOCS, Goodman et al., 1989) to test severity of OCD symptoms; the Obsessive Beliefs Questionnaire (OBQ, developed by the Obsessive Compulsive Cognitions Working Group in 2001) which had a scale called Importance/Control of Thoughts that measured TAF/Thought Suppression; and the Kentucky Inventory of Mindfulness Skills (KIMS, Baer, Smith, & Allen, 2004) to test whether the intervention resulted in participants becoming more mindful. Participants completed these at the first assessment session, at the end of relaxation/start of mindfulness, mid-way through the 6-week mindfulness course, at the end of mindfulness intervention, and at a 2-month follow-up.

# Method in Action

## Sample

In SCEDs, as with other research, it is important to try and control all factors external to the research that might cause a change in the DV (i.e., the confounding factors). In the case of this research, other treatments were the most likely confounds. I decided to include people who were taking specific medication for anxiety (known as selective serotonin uptake inhibitors [SSRIs]) as excluding this group would mean very few people would be eligible to participate. However, to be included participants on medication were required to have been on a stable dose for at least 2 months before the start of the research, for any effects of this to have stabilized. The National Health Service (NHS) ethical review panel that approved the ethical nature of the research

wanted to ensure that I did not deprive participants of an evidence-based treatment by being included in my study, as they could not undergo another psychological therapy at the same time. I therefore chose to recruit participants from a primary care (General Practitioner [GP] level) mental health service because at the time there was no other therapy for OCD available to patients treated in this service (although you could get therapy in specialist secondary care mental health teams).

## Procedure

People who were referred to the primary mental health service with OCD were invited to attend an assessment session with myself where they completed the YBOCS and other study questionnaires to establish that they had at least a moderate level of symptoms (a YBOCS total score of 16 or above and a score of 6 or more on YBOCS items 1 to 3, which indicated specific problems with obsessive intrusive thoughts). They were given information about the research and the individual wording of their daily ratings scales were developed. It was important ethically that participants were given time to consider whether they wanted to participate; I therefore rang them a week later to see whether they wished to do so, at which point they completed consent forms.

Participants kept daily ratings over the next 2 weeks (this was Phase A: Baseline monitoring), before attending the first relaxation session and starting daily relaxation practice. After 2 weeks of relaxation, I checked whether their ratings had stabilized using the C Statistic. If the ratings were found to be unstable, participants undertook further weeks of relaxation; in practice, two participants stabilized after 2 weeks and the third after 3 weeks.

All the sessions in Phase B (Relaxation Control) and Phase C (Mindfulness) were designed to appear similar to reduce the possibility that different formats might affect the DV. I delivered all these sessions myself and developed the 10-min recordings that participants used on a daily basis to practice relaxation or mindfulness between sessions.

## Analysis and Results

No statistical package available calculated the C Statistic (Young, 1941—not to be confused with a different C Statistic by Bamber, 1975). This meant I had to develop a way of calculating this swiftly by hand from the formula in the original mathematical paper. I did this using an excel spreadsheet, where each of the stages of the formula were calculated one stage at a time (e.g., in one part of the formula, I needed to calculate the mean of the daily ratings, subtract this from each individual daily rating, one by one, and then add up the sum of all of these subtractions). Once I had developed a spreadsheet that calculated the full formula at once it was simply a matter of entering the daily ratings by hand and allowing the spreadsheet to calculate the full formula instantly. When establishing if participants' TAF ratings had stabilized at the end of Phase B (Relaxation Control), I had to hand enter data from the rating booklets that participants had brought in during the session, requiring very swift data input, analysis, and interpretation of 14 to 21 data points.

The C Statistic analysis is conducted for each phase of the SCED. The aim is to test for stability in the baseline or control phase (in this case Phase B: Relaxation Control), demonstrated by a non-significant C Statistic. You then move to the intervention phase and test whether the ratings then demonstrate a significant trend (shown by a significant C Statistic). The nature of the effect (e.g., whether the intervention produced an improvement or worsening) is then determined

by visual inspection of graphs of the data; Figure 1 shows a graph of the hypothesized effect of each of the experimental phases.

Figure 1.

Caption: Graph of the hypothesized effect of the experimental phases A (Baseline Monitoring) B (Relaxation Control) C (Mindfulness) on Thought-Action Fusion (TAF), Thought Suppression and Distress daily ratings.

The standardized measures were presented simply as descriptive statistics, charting scale scores for YBOCS, OBQ, and KIMS at the key time points (e.g., change in phase, mid-therapy, end-therapy, follow up) which bookended the time series data. As hypothesized, overall standardized scales measuring OCD-related symptoms showed no clear improvement from baseline to end of relaxation control, but participants showed clear improvements on OCD symptoms and related beliefs at the end of the mindfulness, with two participants maintaining or improving on this at follow-up. As predicted, they also showed improvements in mindfulness skills following the mindfulness intervention (but not the baseline or relaxation control). The time-series data produced by the daily ratings generally followed the pattern of the standardized scales. As hypothesized (and shown in Figure 1), TAF, Thought Suppression and Distress ratings in Phase A (Baseline Monitoring) and Phase B (Relaxation Control) remained generally stable and high, reflecting participant difficulties in these areas. On occasion there appeared to be some minor initial effect of undertaking relaxation (possibly reflecting a placebo effect, as displayed in Figure 1), but which stabilized quickly. Analysis of Phase C (Mindfulness) ratings when compared with stable Phase B (Relaxation Control) ratings found a statistically significant trend.

Inspection of the graphed data indicated that this represented an improvement across all ratings for all three participants.

# Practical Lessons Learnt

## It May Be Hard to Predict Challenges

I discovered that the things I thought might be a problem for the research were different to the challenges that I encountered in reality. I had worried that participants would struggle to complete ratings everyday but in practice there was almost no missing data throughout the whole study. This may have been because people with OCD often show high levels of perfectionism and feelings of responsibility (Rhéaume, Freeston, Dugas, Letarte, & Ladouceur, 1995) which may have resulted in them being particularly diligent in measure completion. I was also concerned that my results section would be sparse when presenting data for only three participants. In actuality, the richness of detail in the data I collected meant that it took considerable work to summarize the data in a meaningful way that could be understood by readers. In future research using a multiple baseline SCED, I would collect time-series data on a single DV, rather than the three daily ratings used in this research.

## Novel Measures Increase the Likelihood of Misunderstanding

One of the three participants who completed the research misunderstood the wording/nature of the Thought Suppression daily rating question: “How much did you try and suppress your obsessive thought and push it completely from your mind?” This participant identified their misunderstanding half-way through the mindfulness intervention, meaning that I had to discount

all of their Thought Suppression data. This is perhaps an inevitable occurrence when the measures have not been through psychometric testing and when relying on single-item measures. A scale composed from a larger number of items would be more robust against the effects of a misunderstanding of the wording of one item.

## Recruitment Can Be Challenging, Even When You Need a Small Number of Participants

I had anticipated recruiting six participants but managed only five, two of whom dropped out—one due to a sudden relocation for work and one because they found it too difficult to do the daily relaxation/mindfulness exercises. In addition, two people with OCD were referred to the research but did not score sufficiently highly on the YBOCS items relating to intrusive thoughts, making them ineligible. On reflection, recruiting from a service who did not normally offer an OCD intervention was going to prove challenging because they were unlikely to receive a large number of referrals for individuals with this type of difficulty. Overall, I learnt that when you need a small number of participants to complete a time-intensive procedure (collect data reliably, every day for around 70 days and complete an intervention task daily for most of this time) you are vulnerable to recruitment difficulties and drop-outs and this arguably has more of an effect when presenting data on a small number of participants compared with drop-outs from studies collecting group data. This emphasizes the need to keep the requirements made of participants to the absolute minimum required to produce quality research.

## Consider the Role of Statistical Analysis in SCED

I devoted great time and effort to undertaking statistical analysis of the data using the C Statistic when it is equivocal as to whether this was necessary. Most SCEDs do not use statistical analysis and instead rely entirely on visual inspection of graphs instead, as many researchers regard this as the best or simplest way of detecting an effect of the intervention and there is no consensus as to which statistical analysis is the preferable alternative as there are many different options (Smith, 2012). However, there are problems with relying entirely on visual inspection, particularly if there is a trend or variability in the baseline (i.e., it is not stable) (Kazdin, 2011) and visual analysis can result in researchers over-estimating the effect of the intervention and can produce inconsistent interpretations of the data (Matyas & Greenwood, 1990). As I would have to justify my analysis in a viva voce examination, I wanted to undertake as robust an analysis as possible so selected both visual inspection and statistical analysis, as recommended by a panel of experts tasked by the What Works Clearinghouse to agree standards for SCEDs (Kratochwill et al., 2010).

# Conclusion

Using a multiple baseline SCED across participants provided the best way to test this new approach to therapy for OCD and provided initial, highly detailed evidence that this intervention could be helpful in reducing OCD symptoms in three individuals. It was particularly useful to see the process of *how* the intervention worked (i.e., by reducing TAF and Thought Suppression), compared with simply discovering whether it reduced symptom levels. Group studies with larger participant numbers are not designed to give you such an insight into the therapy process and as a therapist myself I found it helpful to see the importance of targeting TAF and Thought Suppression in treatment.

The small-N nature of the research was appropriate to the experimental nature of the intervention: you would not first test a new idea, with the potential for being either unhelpful or even harmful, on a large number of people and undertake less thorough monitoring of progress. However, there are clearly considerable limitations to this research that very substantially affect the confidence with which you could conclude that mindfulness would be helpful for other people with OCD. I delivered the relaxation control and mindfulness intervention myself and could not be blinded as to which was the "active" treatment (and indeed, had already hypothesized that mindfulness would be beneficial), so the possibility remains that I unconsciously influenced the results. Equally, as only three people received the full treatment, it is unclear how other people (particularly those with more severe symptoms who would have been seen in a different service to the one from which I recruited) would have responded. Indeed, generalizability is perhaps the primary limitation of SCEDs and other small-N research.

This research was a first step in building the evidence-based for the use of mindfulness with OCD. Since its publication, it has been cited in approximately five academic articles or books a year and the area has now developed to the stage where a protocol for a randomized controlled trial (RCT: the gold-standard research design for testing interventions) into mindfulness-based CBT (MCBT) for OCD has been published (Külz et al., 2014). This protocol cites my research and describes plans to test MCBT with a much larger number of participants with OCD (64 receiving MCBT and 64 receiving a control treatment). It was rewarding to see how even small studies can contribute to the wider process of developing new treatments for mental health problems.

## Exercises and Discussion Questions

1. Under what circumstances might you use a SCED?

2. What are the limitations of SCEDs?

3. What things do you need to consider when developing idiographic measures (measures that have been tailored for the individual)?

4. What are the arguments for relying on visual inspection alone, without statistical testing, when analyzing time-series data in a SCED?

## Further Reading

Hayes, S. C. (1981). Single case experimental design and empirical clinical practice. *Journal of Consulting and Clinical Psychology, 49*, 193-211.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. Oxford, UK: Oxford University Press.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*, 510-550. doi:10.1037/a0029312.

## Web Resources

Commented [1]: AQ1

Single-case designs technical documentation. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

## References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baer, R. A., Smith, G. T., & Allen, K. B. (2004). Assessment of mindfulness by self-report: The Kentucky inventory of mindfulness skills. *Assessment, 11*, 191-206.

Bamber, D. (1975). The area above the ordinal dominance graph and the area below the receiver operating characteristic graph. *Journal of Mathematical Psychology*, *12*, 387-415.

Eisen, J. L., Goodman, W. K., Keller, M. B., Warshaw, M. G., DeMarco, L. M., Luce, D. D., & Rasmussen, S. A (1999). Patterns of remission and relapse in obsessive-compulsive disorder: A 2-year prospective study. *Journal of Clinical Psychiatry*, *60*, 346-351.

Gava, I., Barbui, C., Aguglia, E., Carlino, D., Churchill, R., De Vanna, M., & McGuire, H. (2007). Psychological treatments versus treatment as usual for obsessive compulsive disorder (OCD). *Cochrane Database of Systematic Reviews,* Article CD005333. doi:10.1002/14651858.CD005333.pub2

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischman, R. L., Hill, C. L., . . . Charney, D. S. (1989). The Yale-brown Obsessive Compulsive Scale: I. Development, use and reliability. *Archives of General Psychiatry, 46*, 1006-1011.

Greist, J. H., Marks, I. M., & Baer, L. (2002). Behavior therapy for obsessive–compulsive disorder guided by a computer or by a clinician compared with relaxation as a control. *Journal of Clinical Psychiatry, 63,* 138-145.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and applied settings*. New York, NY: Oxford University Press.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). *Single-case designs technical documentation*. Retrieved on 20.4.2017 from https://ies.ed.gov/ncee/wwc/Document/229

Külz, A. K., Landmann, S., Cludius, B., Hottenrott, B., Rose, N., Heidenreich, T., & Moritz, S. (2014). Mindfulness-based cognitive therapy in obsessive-compulsive disorder: Protocol of a randomized controlled trial. *BMC Psychiatry*, *14*, 314.

Masuda, A., Hayes, S. C., Sackett, C. F., & Twohig, M. P. (2004). Cognitive defusion and self-relevant negative thoughts: Examining the impact of a ninety year old technique. *Behaviour Research and Therapy*, *42*, 477-485.

Matyas, T. A., & Greenwood, K. M. (1990). Visual analysis of single-case time series: Effects of variability, serial dependence, and magnitude of intervention effects. *Journal of Applied Behavior Analysis*, *23*, 341-351.

NICE. (2005). Obsessive-compulsive disorder: Core interventions in the treatment of obsessive-compulsive disorder and body dysmorphic disorder (NICE clinical guideline). Retrieved from http://guidance.nice.org.uk/CG31

Obsessive Compulsive Cognitions Working Group (OCCWG). (2001). Development and initial validation of the obsessive beliefs questionnaire and the interpretation of intrusions inventory. *Behaviour Research and Therapy*, *39*, 987-1006.

Rachman, S. (1993). Obsessions, responsibility and guilt. *Behaviour Research and Therapy, 31*, 149-154.

Rachman, S., & de Silva, P. (1978). Abnormal and normal obsessions. *Behaviour Research and Therapy, 16*, 233-248.

Rhéaume, J., Freeston, M. H., Dugas, M. J., Letarte, H., & Ladouceur, R. (1995). Perfectionism, responsibility and obsessive-compulsive symptoms. *Behaviour Research and Therapy, 33*, 785-794.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of published research and current standards. *Psychological Methods, 17*, 510-550. doi:10.1037/a0029312

Wenzlaff, R. M., & Wegner, D. M. (2000). Thought suppression. *Annual Review of Psychology, 51*, 59-91.

Young, L. C. (1941). On randomness in ordered sequences. *Annals of Mathematical Statistics, 12*, 293-300.